Improved Mixing Condition on the Grid for Counting and Sampling Independent Sets

Ricardo Restrepo* Jinwoo Shin† Prasad Tetali⋆ Eric Vigoda†
Linji Yang†

January 20, 2013

Abstract

The hard-core model has received much attention in the past couple of decades as a lattice gas model with hard constraints in statistical physics, a multicast model of calls in communication networks, and as a weighted independent set problem in combinatorics, probability and theoretical computer science.

In this model, each independent set I in a graph G is weighted proportionally to $\lambda^{|I|}$, for a positive real parameter λ. For large λ, computing the partition function (namely, the normalizing constant which makes the weighting a probability distribution on a finite graph) on graphs of maximum degree $\Delta \geq 3$, is a well known computationally challenging problem. More concretely, let $\lambda_c(\mathbb{T}_\Delta)$ denote the critical value for the so-called uniqueness threshold of the hard-core model on the infinite Δ-regular tree; recent breakthrough results of Dror Weitz (2006) and Allan Sly (2010) have identified $\lambda_c(\mathbb{T}_\Delta)$ as a threshold where the hardness of estimating the above partition function undergoes a computational transition.

We focus on the well-studied particular case of the square lattice $\mathbb{Z}^2$, and provide a new lower bound for the uniqueness threshold, in particular taking it well above $\lambda_c(\mathbb{T}_4)$. Our technique refines and builds on the tree of self-avoiding walks approach of Weitz, resulting in a new technical sufficient criterion (of wider applicability) for establishing strong spatial mixing (and hence uniqueness) for the hard-core model. Our new criterion achieves better bounds on strong spatial mixing when the graph has extra structure, improving upon what can be achieved by just using the maximum degree. Applying our technique to $\mathbb{Z}^2$ we prove that strong spatial mixing holds for all $\lambda < 2.3882$, improving upon the work of Weitz that held for $\lambda < 27/16 = 1.6875$. Our results imply a fully-polynomial *deterministic* approximation algorithm for estimating the partition function, as well as rapid mixing of the associated Glauber dynamics to sample from the hard-core distribution.

*School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332. Email: {restrepo,tetali}@math.gatech.edu. Research supported in part by NSF grant CCF-0910584.

†School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332. Email: {jshin72,ljyang,vigoda}@cc.gatech.edu. Research supported in part by NSF grants CCF-0830298 and CCF-0910584. Jinwoo Shin was supported by the Algorithms and Randomness Center at Georgia Tech.

1 Introduction

In this paper we study phase transitions for sampling weighted independent sets (weighted by an activity $\lambda > 0$) of the 2-dimensional integer lattice $\mathbb{Z}^2$. In statistical physics terminology, we study the hard-core lattice gas model ([6, 13]), which is a simple model of a gas whose particles have non-negligible size (thus preventing them from occupying neighboring sites), with activity $\lambda \in \mathbb{R}_+$ corresponding to the so-called fugacity of the gas. More formally, for a finite graph $G = (V, E)$, let $\Omega = \Omega(G)$ denote the set of independent sets of G. Given an independent set $\sigma \in \Omega$, its weight is defined as $w(\sigma) = \lambda^{|\sigma|}$ and $v \in V$ is said to be occupied under σ if $v \in \sigma$. The associated Gibbs (or Boltzmann) distribution $\mu = \mu_{G,\lambda}$ is defined on Ω as $\mu(\sigma) = w(\sigma)/Z$, where $Z = Z(G, \lambda) = \sum_{\eta \in \Omega} w(\eta)$ is commonly referred to as the partition function.

Recall that Valiant [33] showed that *exactly computing* the number of independent sets is #P-complete, even when restricted to 3-regular graphs (see Greenhill [16]). Hence, we focus our attention on approximation algorithms for estimating the number, or more generally, the partition function. It is well known [17] that the problem of approximating the partition function Z and that of sampling from a distribution that is close to the Gibbs distribution μ, are polynomial-time reducible to each other (see also [31]).

The fundamental notion of a phase transition for a statistical mechanics model on an infinite graph addresses the critical point at which the model starts to exhibit a certain long-range dependence, as a system parameter is varied. In particular, the so-called critical inverse temperature β_c for the Ising or the Potts model, and the critical activity λ_c for the hard-core lattice gas model, are prime examples where the system undergoes a transition from uniqueness to multiplicity of the infinite-volume Gibbs measures.

Phase transition in the hard-core model is also intimately related to the computational complexity of estimating the partition function Z. Recently, a remarkable connection was established between the computational complexity of approximating the partition function for graphs of maximum degree Δ and the phase transition $\lambda_c(\mathbb{T}_\Delta)$ for the infinite regular tree $\mathbb{T}_\Delta$ of degree Δ. On the positive side, Weitz [34] showed a deterministic fully-polynomial time approximation algorithm (FPAS) for approximating the partition function for any graph with maximum degree Δ, when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ and Δ is constant. On the other side, Sly [30] recently showed that for every $\Delta \geq 3$, it is NP-hard (unless NP=RP) to approximate the partition function for graphs of maximum degree Δ, when $\lambda_c(\mathbb{T}_\Delta) < \lambda < \lambda_c(\mathbb{T}_\Delta) + \epsilon_\Delta$, for some function $\epsilon_\Delta > 0$. More recently, Galanis et al. [12] improved the range of λ in Sly's inapproximability result, extending it to all $\lambda > \lambda_c(\mathbb{T}_\Delta)$ for the cases $\Delta = 3$ and $\Delta \geq 6$.

1.1 Prior history and current work

Our work builds upon Weitz's work to get improved results for specific graphs of interest. We focus our attention on what is arguably the simplest, not yet well-understood, case of interest namely the square grid, or the 2-dimensional integer lattice $\mathbb{Z}^2$. Empirical evidence suggests that the critical point $\lambda_c(\mathbb{Z}^2) \approx 3.796$ [13, 3, 26], but rigorous results are significantly far from this conjectured point. The possibility of there being multiple such λ_c is not ruled out, although no one believes that this is the case.

From below, van den Berg and Steif [6] used a disagreement percolation argument to prove $\lambda_c(\mathbb{Z}^2) > \frac{p_c}{1-p_c}$ where p_c is the critical probability for site percolation on $\mathbb{Z}^2$. Applying the best known lower bound on $p_c > 0.556$ for $\mathbb{Z}^2$ by van den Berg and Ermakov [5] implies $\lambda_c(\mathbb{Z}^2) > 1.252\ldots$. Prior to that work, an alternative approach aimed at establishing the Dobrushin-

Shlosman criterion [10], yielded, via computer-assisted proofs, $\lambda_c(\mathbb{Z}^2) > 1.185$ by Radulescu and Styer [28], and $\lambda_c(\mathbb{Z}^2) > 1.508$ by Radulescu [27].

These results were improved upon by Weitz [34] who showed that $\lambda_c(\mathbb{Z}^2) \geq \lambda_c(\mathbb{T}_4) = 27/16 = 1.6875$, where $\mathbb{T}_\Delta$ is the infinite, complete, regular tree of degree Δ. For the upper bound, a classical Peierls' type argument implies $\lambda_c(\mathbb{Z}^2) = O(1)$ [9]. (A related result of Randall [29] showing slow mixing of the Glauber dynamics for $\lambda > 8.066$ gives hope for a better upper bound on $\lambda_c(\mathbb{Z}^2)$.) The regular tree $\mathbb{T}_\Delta$ is one of the only examples (that we know of) where the critical point is known exactly, and in this case, Kelly [18] showed that $\lambda_c(\mathbb{T}_\Delta) = (\Delta-1)^{\Delta-1}/(\Delta-2)^\Delta$.

In this work we present a new general approach which, for the case of the hard-core model on $\mathbb{Z}^2$, improves the lower bound to $\lambda_c(\mathbb{Z}^2) > 2.3882$. There are various algorithmic implications for finite subgraphs of the $\mathbb{Z}^2$ when $\lambda < 2.3882$. Our results imply that Weitz's deterministic FPAS is also valid on subgraphs of $\mathbb{Z}^2$ for the same range of λ. Thanks to the existing literature on general spin systems ([22, 23, 8, 11]), our results also imply that the Glauber dynamics has $O(n \log n)$ mixing time for any finite subregion $G = (V, E)$ of $\mathbb{Z}^2$ when $\lambda < 2.3882$, where $n = |V|$. Recall that the Glauber dynamics is a simple Markov chain that updates the configuration at a randomly chosen vertex in each step, see [19] for an introduction to the Glauber dynamics. The stationary distribution of this chain is the Gibbs distribution. Hence, it is of interest as an algorithmic technique to randomly sample from the Gibbs distribution, and also as a model of how physical systems reach equilibrium. The mixing time is the number of steps (from the worst initial configuration) until the distribution is guaranteed to be within variation distance $\leq 1/4$ of the stationary distribution.

As in Weitz's work, our approach can be used for other 2-spin systems, such as the Ising model. This is discussed in Section 6. As will be evident from the following high-level idea of our approach, it can be applied to other graphs of interest. Our work also provides an arguably simpler way to derive the main technical result of Weitz showing that any graph with maximum degree Δ has strong spatial mixing (SSM) when $\lambda < \lambda_c(\mathbb{T}_\Delta)$.

To underline the difficulty in estimating bounds on λ_c, we remark that the *existence* of a (unique) critical activity λ_c remains conjectural and an open problem for $\mathbb{Z}^d$, for $d \geq 2$. In contrast, for the Ising model, the critical inverse temperature $\beta_c(\mathbb{Z}^2)$ has been known since 1944 [24]; interestingly, the corresponding critical point for the q-state Potts model (for $q \geq 2$) has only recently been established (by Beffara and Duminil-Copin [4]) to be $\beta_c(q) = \log(1 + \sqrt{q})$, settling a long-standing open problem. The lack of monotonicity in λ in the hard-core model poses a serious challenge in establishing such a sharp result for this model. In fact, Brightwell et al. [7] showed that in general such a monotonicity need not hold, by providing an example with a non-regular tree.

2 Technical Preliminaries and Proof Approach

Before presenting our approach, it is useful to review briefly the uniqueness/non-uniqueness phase transition, and introduce associated notions of decay of spatial correlation, known as weak and strong spatial mixing properties. Much of the below discussion is simplified for the case of the hard-core model on $\mathbb{Z}^2$, wherein one utilizes certain induced monotonicity (given by the bipartite property) in the model and the amenability of the graph.

2.1 Uniqueness, Weak and Strong Spatial Mixing

Let B_L denote the finite graph corresponding to a box of side-length $2L+1$ centered around the origin in $\mathbb{Z}^2$. Thus, $B_L = (V, E)$, where $V = (i, j) \in \mathbb{Z}^2 : -2L-1 \le i, j \le 2L+1$ with edges between pairs of vertices at L_1 distance (or Manhattan distance) equal to one. Since this is a bipartite graph, we may fix one such partition $V = \text{even} \cup \text{odd}$ – for example, it is standard to consider the set of vertices at an even distance from the origin as the even set. The boundary of B_L are those vertices $v = (v_1, v_2) \in V$ where $|v_i| = 2L+1$ for $i = 1$ or $i = 2$. The hard-core model on bipartite graphs is a monotone system (e.g., see [11]), which for the current discussion implies that we only have to consider two assignments to the boundary: all even vertices or all odd vertices on the boundary are occupied. Let $\alpha^{\text{even}}_{L,r}$ ($\alpha^{\text{odd}}_{L,r}$) denote the marginal probability that the origin r is unoccupied given the even (odd, respectively) boundary. Then to establish uniqueness of the Gibbs measures, we need that:

$$\lim_{L\to\infty} |\alpha^{\text{even}}_{L,r} - \alpha^{\text{odd}}_{L,r}| = 0.$$

We are interested in the critical point λ_c for the transition between uniqueness and non-uniqueness. A standard way to establish uniqueness is by proving one of the spatial mixing properties introduced next.

Let $G = (V, E)$ be a (finite) graph. For $S \subset V$, a configuration $\boldsymbol{\rho}$ on S specifies a subset of S as occupied and the remainder as unoccupied. Let $\mu^{\boldsymbol{\rho}} = \mu^{\boldsymbol{\rho}}_G$ denote the Gibbs distribution conditional on configuration $\boldsymbol{\rho}$ to S. For $v \in V$, let $\alpha^{\boldsymbol{\rho}}_v = \alpha^{\boldsymbol{\rho}}_{G,v}$ denote the marginal probability that v is unoccupied in $\mu^{\boldsymbol{\rho}}$.

The first spatial mixing property is *Weak Spatial Mixing (WSM)*. Here we consider a pair of boundary configurations on a subset S and consider the "influence" on the marginal probability that a vertex v is unoccupied. WSM says that the influence on v decays exponentially in the distance of S from v.

Definition 1 (Weak Spatial Mixing)**.** *For the hard-core model at activity λ, for finite graph $G = (V, E)$,* WSM *holds with rate $\gamma \in (0, 1)$ if for every $v \in V$, every $S \subset V$, and every two configurations $\boldsymbol{\rho}, \boldsymbol{\eta}$ on S,*

$$|\alpha^{\boldsymbol{\rho}}_v - \alpha^{\boldsymbol{\eta}}_v| \;\le\; \gamma^{\operatorname{dist}(v,T)}$$

where $\operatorname{dist}(v, S)$ *is the graph distance (i.e., length of the shortest path) between v and (the nearest point in) the subset S.*

The second property of interest is *Strong Spatial Mixing (SSM)*. The intuition is that if a pair of boundary configurations on a subset S agree at some vertices in S then those vertices "encourage" v to agree. Therefore, SSM says that the influence on v decays exponentially in the distance of v from the subset of vertices where the pair of configurations differ.

Definition 2 (Strong Spatial Mixing)**.** *For the hard-core model at activity λ, for finite graph $G = (V, E)$,* SSM *holds with rate $\gamma \in (0, 1)$ if for every $v \in V$, every $S \subset V$, every $S' \subset S$, and every two configurations $\boldsymbol{\rho}, \boldsymbol{\eta}$ on S where $\boldsymbol{\rho}(S \setminus S') = \boldsymbol{\eta}(S \setminus S')$,*

$$|\alpha^{\boldsymbol{\rho}}_v - \alpha^{\boldsymbol{\eta}}_v| \;\le\; \gamma^{\operatorname{dist}(v,S')}.$$

Note that since $\operatorname{dist}(v, T) \le \operatorname{dist}(v, T \setminus S)$, SSM implies WSM for the same rate. Moreover, it is a standard fact that such an exponential decay in finite boxes (say), in $\mathbb{Z}^d$, implies uniqueness of the corresponding infinite volume Gibbs measure on $\mathbb{Z}^d$, see Georgii [14] for an introduction

to the theory of infinite-volume Gibbs measures. We can specialize the above notions of WSM and SSM to a particular vertex v, in which case we say that WSM or SSM holds *at* v. If the graph is a rooted tree, we will always assume that the notions of WSM and SSM are considered at the root.

For the hard-core model on a graph $G = (V, E)$, for a subset of vertices S and a fixed configuration ρ on S, it is equivalent to consider the subgraph G' which we obtain for each $v \in S$ that is fixed to be unoccupied we remove v from G, and for each $v \in S$ that is fixed to be occupied we remove v and its neighbors $N(v)$ from G. In this way we obtain the following observation which will be useful for proving SSM holds.

Observation 1. *For a graph $G = (V, E)$ and $v \in V$, SSM holds in G at vertex v iff WSM holds for all subgraphs G' (of G) at vertex v. To be precise, by subgraphs we mean graphs obtained by considering all subgraphs of G and taking the component containing v.*

2.2 Self-Avoiding Walk Tree Representation

Since our work builds on that of Weitz's, we first describe the self-avoiding walk (SAW) tree representation introduced in [34]. Given $G = (V, E)$, we first fix an arbitrary ordering $>_w$ on the neighbors of each vertex w in G. For each $v \in V$, the tree $T_{\text{saw}}(G, v)$ is constructed as follows. Consider the tree T of self-avoiding walks originating from v, additionally including the vertices closing a cycle as leaves of the tree. We then fix such leaves of T to be occupied or unoccupied in the following manner. If a leaf vertex closes a cycle in G, say $w \to v_1 \to \dots v_\ell \to w$, then if $v_1 >_w v_\ell$ we fix this leaf to be unoccupied, otherwise if $v_1 <_w v_\ell$ we fix the leaf to be occupied. Note, if the leaf is fixed to be unoccupied we simply remove that vertex from the tree. If the leaf is fixed to be occupied, we remove that leaf and all of its neighbors, i.e. we remove the parent of that leaf from the tree. The resulting tree is denoted as $T_{\text{saw}} = T_{\text{saw}}(G, v)$. See Figure 1 for an illustration of T_{saw} for a particular example.

Weitz [34] proves the following theorem for the hard-core model, which shows that the marginal distribution at the root in $T_{\text{saw}}(G, v)$ is identical to the marginal distribution for v in G. For a graph $G = (V, E)$, a subset $S \subset V$ and configuration ρ on S, for $T_{\text{saw}} = T_{\text{saw}}(G, v)$, let ρ in T_{saw} denote the configuration on S in T_{saw} where for $w \in S$ every occurrence of w in T_{saw} is assigned according to ρ.

Theorem 1 (SAW Tree Representation, Theorem 3.1 in [34])**.** *For any graph $G = (V, E)$, $v \in V$, $\lambda > 0$, and configuration $\boldsymbol{\rho}$ on $S \subset V$, for $T = T_{\text{saw}}(G, v)$ the following holds:*

$$\alpha^{\boldsymbol{\rho}}_{G,v} = \alpha^{\boldsymbol{\rho}}_{T,v}.$$

Note, the tree $T_{\text{saw}}(G, v)$ preserves the distance of vertices from v in G, which implies the following corollary.

Corollary 2. *If SSM holds with rate γ for $T_{\text{saw}}(G, v)$ for all v, then SSM holds for G with rate γ.*

The reverse implication of Corollary 2 does not hold since there are configurations on S in T_{saw} which are not necessarily realizable in G. Observe that if G has maximum degree Δ, any SAW tree of G is a subtree of the regular tree of degree Δ.

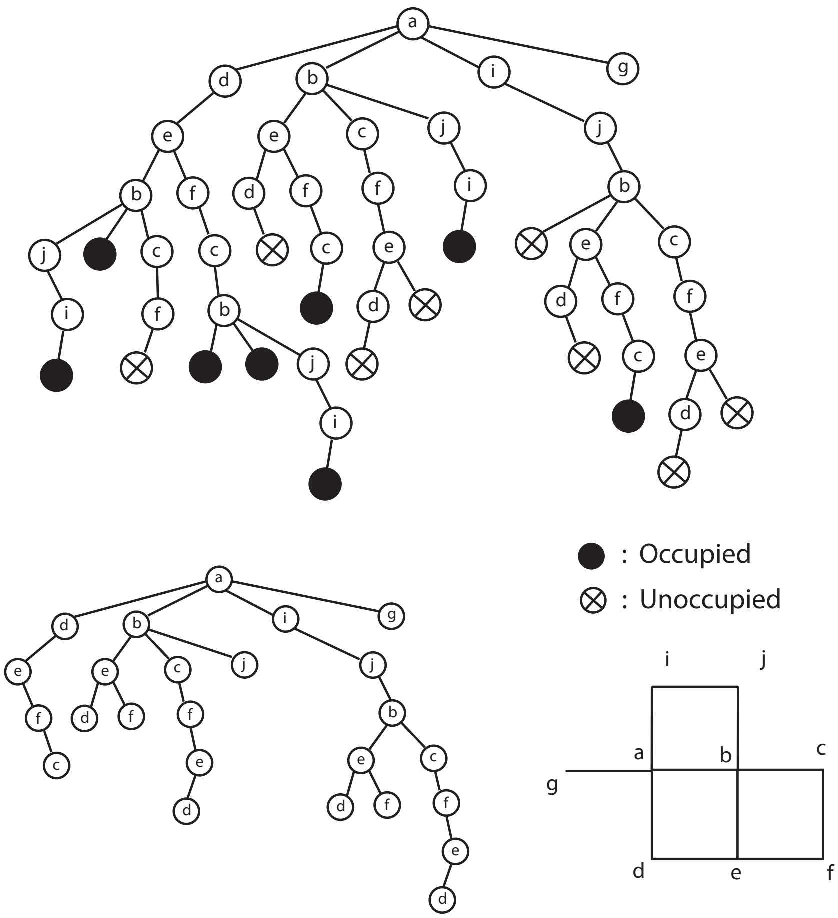

Figure 1: Example of self-avoiding walk tree T_{saw}. The above tree describes $T_{\text{saw}}(G, a)$ with occupied and unoccupied leaves, while the below one is the same tree after removing those assigned leaves. At each vertex, we consider the ordering $N > E > S > W$ of its neighbors where N, E, S, W represent the neighbors in the North, East, South, West directions, respectively.

2.3 Our Proof Approach

In summary, Weitz [34] first shows (via Theorem 1) that to prove SSM holds on a graph $G = (V, E)$, it suffices to prove SSM holds on the trees $T_{\text{saw}}(G, v)$, for all $v \in V$. Weitz then proves that the regular tree $\mathbb{T}_\Delta$ "dominates" every tree of maximum degree Δ in the sense that, for all trees of maximum degree Δ, SSM holds when $\lambda < \lambda_c(\mathbb{T}_\Delta)$. We refine this second part of Weitz's approach. In particular, for graphs with extra structure, such as $G = \mathbb{Z}^2$, we bound $T_{\text{saw}}(\mathbb{Z}^2)$ by a tree T^* that is much closer to it than the regular tree $\mathbb{T}_\Delta$. We then establish a criterion that achieves better bounds on SSM for trees when the trees have extra structure.

The tree T^* will be constructed in a regular manner so that we can prove properties about it – the construction of T^* is governed by a (progeny) $t \times t$ matrix $\boldsymbol{M}$, whose rows correspond to t types of vertices, with the entry M_{ij} specifying the number of children of type j that a vertex of type i begets. We will then show a sufficient condition using entries of $\boldsymbol{M}$ which implies that SSM holds for T^* and for any subgraph of T^*, including $T_{\text{saw}}(\mathbb{Z}^2)$. The construction of T^* is reminiscent of the strategy employed in [1, 25] to upper bound the connectivity constant

of several lattice graphs, including $\mathbb{Z}^2$. The derivation of our sufficient condition has some inspiration from belief propagation algorithms.

As a byproduct of our proof that our new criterion implies SSM for T^*, we get a new (and simpler) proof of the second part of Weitz's approach, namely, that for all trees of maximum degree Δ, SSM holds when $\lambda < \lambda_c(\mathbb{T}_\Delta)$.

3 Branching Matrices and Strong Spatial Mixing

As alluded to above, we will utilize more structural properties of self-avoiding walk trees. To this end, we consider families of trees which can be recursively generated by certain rules; we then show that such a general family is also analytically tractable.

3.1 Definition of Branching Matrices

We say that the matrix $\boldsymbol{M}$ is a $t \times t$ branching matrix if every entry M_{ij} is a non-negative integer. We say the maximum degree of $\boldsymbol{M}$ is $\Delta = \Delta(\boldsymbol{M}) = \max_{1\le i\le t} \sum_{1\le j\le t} M_{ij}$, the maximum row sum. Given a branching matrix $\boldsymbol{M}$, we define the following family of graphs. In essence, it includes a graph G if the self-avoiding walk trees of G can be generated by $\boldsymbol{M}$.

Definition 3 (Branching Family)**.** *Given a $t \times t$ branching matrix $\boldsymbol{M}$, $\mathcal{F}_{\le \boldsymbol{M}}$ includes trees which can be generated under the following restrictions:*

- *Each vertex in tree $T \in \mathcal{F}_{\le \boldsymbol{M}}$ has its type $i \in \{1, \dots, t\}$.*
- *Each vertex of type i has at most M_{ij} children of type j.*

In addition, we use the notation $G = (V, E) \in \mathcal{F}_{\le \boldsymbol{M}}$ if $T_{\text{saw}}(G, v) \in \mathcal{F}_{\le \boldsymbol{M}}$ for all $v \in V$.

For example, the family $\mathcal{F}_{\le \boldsymbol{M}}$ with $\boldsymbol{M} = [\Delta]$ includes the family of trees with maximum branching Δ. On the other hand, $\mathcal{F}_{\le \boldsymbol{M}}$ with $\boldsymbol{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ describes the family of graphs of maximum degree $\Delta + 1$, by assigning the root of tree $T \in \mathcal{F}_{\le \boldsymbol{M}}$ to be of type 1 and the other vertices of the tree to be of type 2. Note that if $\boldsymbol{M}$ has maximum degree Δ, then every $G \in \mathcal{F}_{\le \boldsymbol{M}}$ also has maximum degree Δ.

In this framework, Weitz's result establishing SSM for all graphs of maximum degree Δ when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ can be stated as establishing SSM with uniform rate for all $G \in \mathcal{F}_{\le \boldsymbol{M}}$ with $\boldsymbol{M} = \begin{pmatrix} 0 & \Delta \\ 0 & \Delta - 1 \end{pmatrix}$; and we are interested in establishing its analogy for general $\boldsymbol{M}$. To this end, we will use the following notion of SSM for $\boldsymbol{M}$.

Remark 1. *To establish SSM for $\boldsymbol{M}$, it suffices to prove that SSM holds with uniform rate for all trees in $\mathcal{F}_{\le \boldsymbol{M}}$ due to Corollary 2. In addition, note that SSM holds for $\boldsymbol{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ if and only if it holds for (Δ) since the root of a tree $T \in \mathcal{F}_{\le \boldsymbol{M}}$ is the only possible vertex of type 1 in T.*

Finally, we define SSM for a branching matrix $\boldsymbol{M}$.

Definition 4. *Given a branching matrix $\boldsymbol{M}$, we say SSM holds for $\boldsymbol{M}$ if SSM holds with uniform rate for all $G \in \mathcal{F}_{\le \boldsymbol{M}}$.*

Remark 2. *To establish SSM for $\boldsymbol{M}$, it suffices to prove that SSM holds with uniform rate for all trees in $\mathcal{F}_{\leq \boldsymbol{M}}$ due to Corollary 2.*

3.2 Implications of SSM

We present a new approach for proving SSM for a branching matrix $\boldsymbol{M}$. There are multiple consequences of SSM for $\boldsymbol{M}$ as summarized in the following theorem. We first state some definitions needed for stating the theorem.

Following Goldberg et al. [15] we use the following variant of amenability for infinite graphs. Here we consider an infinite graph $G = (V, E)$. For $v \in V$ and a non-negative integer d, let $\mathrm{B}_d(v)$ denote the set of vertices within distance $\leq d$ from v, where distance is the length of the shortest path. For a set of vertices S, the (outer) boundary and neighborhood amenability are defined, respectively, as:

$$\partial S := \{w \in V : w \notin S, \text{ and } w \text{ has a neighbor } y \in S\} \quad \text{and} \quad r_d = \sup_{v \in V} \frac{|\partial \mathrm{B}_d(v)|}{|\mathrm{B}_d(v)|}.$$

The infinite graph is said to be *neighborhood-amenable* if $\inf_d r_d = 0$.

Now we can state the following theorem detailing the implications of SSM of interest to us.

Theorem 3. *For a $t \times t$ branching matrix $\boldsymbol{M}$, if SSM holds for $\boldsymbol{M}$ then the following hold:*

1. *For every $G \in \mathcal{F}_{\leq \boldsymbol{M}}$, SSM holds on G.*
2. *For every infinite graph $G \in \mathcal{F}_{\leq \boldsymbol{M}}$, there is a unique infinite-volume Gibbs measure on G.*
3. *If $\boldsymbol{M}$ has maximum degree Δ, if $t = O(1)$ and $\Delta = O(1)$, then for every (finite) $G \in \mathcal{F}_{\leq \boldsymbol{M}}$, Weitz's algorithm [34] gives an FPAS for approximating the partition function $Z(G)$.*
4. *For every infinite $H \in \mathcal{F}_{\leq \boldsymbol{M}}$ which is neighborhood-amenable, for every finite subgraph $G = (V, E)$ of H, the Glauber dynamics has $O(n^2)$ mixing time. Moreover, if $H = \mathbb{Z}^d$ for constant d, then for every finite subgraph $G = (V, E)$ of H, the Glauber dynamics has $O(n \log n)$ mixing time.*

Proof. Part 1 is by the definition of SSM for $\boldsymbol{M}$. The uniqueness result follows from the fact that the infinite-volume extremal Gibbs measures on the infinite graph G can be obtained by taking limits of finite measures, see Georgii [14] for an introduction to infinite-volume Gibbs measures, and see Martinelli [21] for Part 2. Part 3 immediately follows from the work of Weitz [34]. Finally, for Part 4, there is a long line of work showing that for the integer lattice $\mathbb{Z}^d$ in fixed dimensions, for the Ising model SSM on $\mathbb{Z}^d$ implies $O(n \log n)$ mixing time of the Glauber dynamics on finite subregions of $\mathbb{Z}^d$, e.g., see Cesi [8] and Martinelli [21] (and the references therein) for recent results on this problem. These results for the Ising model are typically stated for a general class of models, but that class does not include models with hard constraints, such as the hard-core model studied here. Dyer et al. [11] showed a simpler proof for the hard-core model that utilizes the monotonicity of the model. We use this result of [11] in Theorem 8 to get $O(n \log n)$ mixing time for subregions of $\mathbb{Z}^2$. Goldberg et al. [15, Theorem 8] showed that for k-colorings, if SSM holds for an infinite graph G that is neighborhood-amenable, the Glauber dynamics has $O(n^2)$ mixing time for all finite subgraphs of G. Their proof holds for the hard-core model which implies Part 4. $\square$

4 Establishing SSM for Branching Matrices

In this section we present a sufficient condition implying SSM for the family of trees generated by a branching matrix. As a consequence of the approach presented in this section we get a simpler proof of Weitz's result [34] implying SSM for all graphs with maximum degree Δ when $\lambda < \lambda_c(\mathbb{T}_\Delta)$. We then apply the condition presented in this section to $\mathbb{Z}^2$ in Section 5.

To show the decay of influence of a boundary condition $\boldsymbol{\rho}$, a common strategy is to prove some form of contraction for the 'one-step' iteration given in (1) below. More generally, we will prove such a contraction for an appropriate set of 'statistics' of the unoccupied marginal probability.

A *statistic* of the univariate parameter $x \in [a, b]$ is a monotone (i.e., strictly increasing or decreasing) function $\varphi : [a, b] \to \mathbb{R}$. For a $t \times t$ branching matrix $\boldsymbol{M}$ we consider a set of t statistics $\varphi_1, \ldots, \varphi_t$, one for each type. For the simpler case when $\boldsymbol{M} = [\Delta]$ and hence $t = 1$, we have a single statistic φ. Our aim is proving contraction for an appropriate set of statistics of the probability that the root of a tree is unoccupied.

We first focus on the case of a single type. Consider a tree $T = (V, E) \in \mathcal{F}_{\leq \boldsymbol{M}}$ with root r. For $v \in V$, let $N(v)$ denote the children of v, and let $d(v) := |N(v)|$ the number of children. Let T_v denote the subtree rooted at v. We will analyze the unoccupied probability for a vertex v, but v will always be the root of its subtree. Hence, to simplify the notation, for a boundary condition $\boldsymbol{\rho}$ on $S \subset V$, let $\alpha_v^{\boldsymbol{\rho}} = \alpha_{T_v,v}^{\boldsymbol{\rho}}$.

A straightforward recursive calculation with the partition function leads to the following relation:

$$\alpha_v^{\boldsymbol{\rho}} = \begin{cases} \frac{1}{1+\lambda} & \text{if } N(v) = \emptyset \\ \frac{1}{1+\lambda \prod_{w \in N(v)} \alpha_w^{\boldsymbol{\rho}}} & \text{otherwise.} \end{cases} \tag{1}$$

Note, the unoccupied probability always lies in the interval $I := \left[\frac{1}{1+\lambda}, 1\right]$, i.e., for all v, all ρ, $\alpha_v^\rho \in I$.

For $v \in V$, let $m_v^{\boldsymbol{\rho}} := \varphi(\alpha_v^{\boldsymbol{\rho}})$ be the 'message' at vertex v. The messages satisfy the following recurrence:

$$m_v^{\boldsymbol{\rho}} = \varphi\left(\frac{1}{1+\lambda \prod_{w \in N(v)} \alpha_w^{\boldsymbol{\rho}}}\right) = \varphi\left(\frac{1}{1+\lambda \prod_{w \in N(v)} \varphi^{-1}(m_w^{\boldsymbol{\rho}})}\right).$$

Our aim is to prove uniform contraction of the messages on all trees $T \in \mathcal{F}_{\leq \boldsymbol{M}}$. To this end, we will consider a more general set of messages. Namely, we consider messages $m_1, \ldots, m_\Delta$ where for every $1 \leq i \leq \Delta$, $m_i = \varphi(\alpha_i)$ and $\alpha_i \in I := \left[\frac{1}{1+\lambda}, 1\right]$. This set of tuples $\alpha_1, \ldots, \alpha_\Delta \in I$ contains all of the tuples obtainable on a tree.

For $\alpha_1, \ldots, \alpha_\Delta \in I$, let $m_i = \varphi(\alpha_i), 1 \leq i \leq \Delta$, and let

$$F(m_1, \ldots, m_\Delta) := \varphi\left(\frac{1}{1+\lambda \prod_{i=1}^{\Delta} \varphi^{-1}(m_i)}\right).$$

Ideally, we would like to establish the following contraction: there exists a $0 < \gamma < 1$ such that for all $\alpha_1, \ldots, \alpha_\Delta, \alpha'_1, \ldots, \alpha'_\Delta \in I$,

$$|F(m_1, \ldots, m_\Delta) - F(m'_1, \ldots, m'_\Delta)| \leq \gamma \max_{1 \leq i \leq \Delta} |m_i - m'_i|,$$

where $m_i = \varphi(\alpha_i)$ and $m'_i = \varphi(\alpha'_i)$. We will instead show that the following weaker condition suffices. Namely, that the desired contraction holds for all $|\alpha_i - \alpha'_i| \leq \epsilon$ for some $\epsilon > 0$. This is equivalent to the following condition.

Definition 5. *Let $I = \left[\frac{1}{1+\lambda}, 1\right]$. For the branching matrix $\boldsymbol{M} = [\Delta]$, we say that* Condition ($\star$) *is satisfied if for all $\alpha_1, \ldots, \alpha_\Delta \in I$, by setting $m_i = \varphi(\alpha_i)$ for $1 \le i \le \Delta$, the following holds:*

$$\|\nabla F(m_1, \ldots, m_\Delta)\|_1 = \sum_{i=1}^{\Delta} \left| \frac{\partial F(m_1, \ldots, m_\Delta)}{\partial m_i} \right| < 1. \tag{$\star$}$$

Let us now consider a natural generalization of the above notion for a branching matrix with multiple types. Let $\boldsymbol{M}$ be a $t \times t$ branching matrix. For $1 \le \ell \le t$, let $\Delta_\ell = \sum_{k=1}^{t} M_{\ell k}$ denote the maximum number of children of a vertex of type ℓ. Once again, consider a tree $T = (V, E) \in \mathcal{F}_{\le \boldsymbol{M}}$ with root r. For $v \in V$, let $t(v)$ denote its type. As before, $N(v)$ are the children of v, $d(v)$ is the number of children of v, and for a boundary condition $\boldsymbol{\rho}$ on $S \subset V$, $\alpha_v^{\boldsymbol{\rho}}$ is the unoccupied probability for v in the tree T_v under ρ.

The recursive calculation in (1) for α_v in terms of $\alpha_w, w \in N(v)$, still holds. For the case of multiple types, for $v \in V$, let $m_v^{\boldsymbol{\rho}} := \varphi_{t(v)}(\alpha_v^{\boldsymbol{\rho}})$ be the message at vertex v. The messages satisfy the following recurrence:

$$m_v^{\boldsymbol{\rho}} = \varphi_{t(v)}\left(\frac{1}{1 + \lambda \prod_{w \in N(v)} \varphi_{t(w)}^{-1}(m_w^{\boldsymbol{\rho}})} \right).$$

For each type $1 \le \ell \le t$, we consider contraction of messages derived from all $\alpha_1, \ldots, \alpha_{\Delta_\ell} \in I$. We need to identify the type of each these quantities α_i in order to determine the appropriate statistic to apply. The assignment of types needs to be consistent with the branching matrix $\boldsymbol{M}$. Hence, let $s_\ell : \{1, \ldots, \Delta_\ell\} \to \{1, \ldots, t\}$ be the following assignment. Let $M_{\ell, \le 0} = 0$ and for $1 \le i \le t$, let $M_{\ell, \le i} = \sum_{k=1}^{i} M_{\ell, k}$. For $1 \le i \le t$, for $M_{\ell, \le i-1} < j \le M_{\ell, \le i}$, let $s_\ell(j) = i$.

For type $1 \le \ell \le t$, for $\alpha_1, \ldots, \alpha_{\Delta_\ell} \in I$, set $m_j = \varphi_{s_\ell(j)}(\alpha_j), 1 \le j \le \Delta_\ell$, and let

$$F_\ell(m_1, \ldots, m_{\Delta_\ell}) := \varphi_\ell \left(\frac{1}{1 + \lambda \prod_{j=1}^{\Delta_\ell} \varphi_{s_\ell(j)}^{-1}(m_j)} \right).$$

Note,

$$m_v^{\boldsymbol{\rho}} = F_{t(v)}\left(m_{w_1}^{\boldsymbol{\rho}}, \ldots, m_{w_{d(v)}}^{\boldsymbol{\rho}}\right) \quad \text{where} \quad N(v) = \{w_1, \ldots, w_{d(v)}\}.^{1} \tag{2}$$

We generalize Condition ($\star$) to branching matrices with multiple types by allowing a weighting of the types by parameters $c_1, \ldots, c_t$.

Definition 6. *Let $I = \left[\frac{1}{1+\lambda}, 1\right]$. For a $t \times t$ branching matrix $\boldsymbol{M}$, we say that* Condition ($\star\star$) *is satisfied if there exist $c_1, \ldots, c_t$, such that for all $1 \le \ell \le t$, for all $\alpha_1, \ldots, \alpha_{\Delta_\ell} \in I$, by setting $m_i = \varphi_{s_\ell(i)}(\alpha_i)$ for $1 \le i \le \Delta_\ell$, the following holds:*

$$\sum_{i=1}^{\Delta_\ell} c_{s_\ell(i)} \left| \frac{\partial F_\ell(m_1, \ldots, m_{\Delta_\ell})}{\partial m_i} \right| < c_\ell. \tag{$\star\star$}$$

The following lemma establishes a sufficient condition so that SSM holds for $\boldsymbol{M}$.

[1]Strictly speaking, F_ℓ requires Δ_ℓ arguments, so for (2) to hold in the case when $d(v) < \Delta_\ell$ we can simply add additional arguments corresponding to $\alpha = 1$, which fixes these additional vertices to be unoccupied (and therefore absent).

Lemma 4. *For a $t \times t$ branching matrix $\boldsymbol{M}$, if for every $1 \le \ell \le t$, φ_ℓ is continuously differentiable on the interval $I = \left[\frac{1}{1+\lambda}, 1\right]$ and $\inf_{x\in I} |\varphi'_\ell(x)| > 0$, and if Condition ($\star$) is satisfied for $t = 1$ or Condition ($\star\star$) is satisfied for $t \ge 2$ then SSM holds for $\boldsymbol{M}$, and hence the conclusions of Theorem 3 follow.*

Proof. For a tree $T = (V, E)$ with root r, let $\alpha^+_{L,r}$ and $\alpha^-_{L,r}$ denote the marginal probabilities that the root of T is unoccupied conditional on the vertices at level L (i.e., distance L from the root) being occupied and unoccupied, respectively.

The main result for proving Lemma 4 is that there exist $\gamma < 1$ and $L_0 < \infty$ such that for every tree $T \in \mathcal{F}_{\le \boldsymbol{M}}$ and every integer $L \ge L_0$,

$$\left|\alpha^+_{L,r} - \alpha^-_{L,r}\right| \;\le\; \gamma^L. \tag{3}$$

We first explain why (3) implies Lemma 4 and then we prove (3). Consider a tree $T = (V, E)$ with root r, and a boundary condition $\boldsymbol{\rho}$ on $S \subset V$. Set $L = \mathrm{dist}(r, S)$ as the distance of S to the root of T. The hard-core model on bipartite graphs has a monotonicity of boundary conditions (c.f., [11]) which implies that for odd L, $\alpha^+_{L,r} \ge \alpha^{\boldsymbol{\rho}}_r \ge \alpha^-_{L,r}$, and for even L, $\alpha^+_{L,r} \ge \alpha^{\boldsymbol{\rho}}_r \ge \alpha^-_{L,r}$. Hence, for any pair of boundary conditions $\boldsymbol{\rho}$ and $\boldsymbol{\eta}$ on S,

$$|\alpha^{\boldsymbol{\rho}}_r - \alpha^{\boldsymbol{\eta}}_r| \le \left|\alpha^+_{L,r} - \alpha^-_{L,r}\right|.$$

Therefore, by the definition of WSM in Definition 1, proving (3) implies WSM for T. Since this holds for all $T' \in \mathcal{F}_{\le \boldsymbol{M}}$, by Observation 1, it implies SSM for all $T' \in \mathcal{F}_{\le \boldsymbol{M}}$, which, by Remark 2, implies SSM for $\boldsymbol{M}$.

We now turn our attention to proving (3). Fix a $t \times t$ branching matrix $\boldsymbol{M}$ and consider a tree $T = (V, E) \in \mathcal{F}_{\le \boldsymbol{M}}$ with root r. Given $y \in [0, 1]$, let $\beta_{L,v}(y)$ denote the marginal probability that the root of T_v is unoccupied given all of the vertices at level L (in T_v) are assigned marginal probability y of being unoccupied (conditional on its parent being unoccupied). Intuitively, $\beta_{L,v}(y)$ can be thought as the marginal probability conditioned on a 'fractional' boundary configuration at level L. As in (1), $\beta_{L,r}(y)$ satisfies the following recurrence for $y \in [0, 1]$:

$$\beta_{L,r}(y) = \begin{cases} y & \text{if } L = 0, \\ \frac{1}{1+\lambda} & \text{if } L > 0 \text{ and } N(r) = \emptyset, \\ \frac{1}{1+\lambda \prod_{w \in N(r)} \beta_{L-1,w}(y)} & \text{otherwise.} \end{cases} \tag{4}$$

From (4) and (1), it follows that $\alpha^+_{L,r} = \beta_{L,r}(1)$ and $\alpha^-_{L,r} = \beta_{L,r}(0)$. Hence, in order to analyze the messages for $\alpha^+_{L,r}$ and $\alpha^-_{L,r}$, we will analyze the messages for $\beta_{L,r}(y)$. Therefore, for $v \in V$, let $m_{L,v}(y) = \varphi_{t(v)}(\beta_{L,v}(y))$. Analogous to (2), we now have that:

$$m_{L,r}(y) = F_{t(r)}\left(m_{L-1,w_1}(y), \ldots, m_{L-1,w_{d(r)}}(y)\right) \qquad \text{where} \quad N(r) = \{w_1, \ldots, w_{d(r)}\}.$$

Observe that for all $y \in [0, 1]$, all $L > 0$, all $v \in V$, $\beta_{L,v}(y) \in I = \left[\frac{1}{1+\lambda}, 1\right]$, and hence we can use Condition ($\star\star$) to analyze $m_{L,r}$.

Using the fact that $\beta_{L,v}(y)$ and $m_{L,v}(y)$ are continuously differentiable for $y \in [0, 1]$, we have that for $L > 0$,

$$\left|\alpha^+_{L,r} - \alpha^-_{L,r}\right| \;=\; |\beta_{L,r}(1) - \beta_{L,r}(0)| \;\le\; \int_0^1 \left|\frac{\partial \beta_{L,r}(y)}{\partial y}\right| dy \;\le\; \frac{\int_0^1 \left|\frac{\partial m_{L,r}(y)}{\partial y}\right| dy}{\inf_{x\in I}\left|\varphi'_{t(r)}(x)\right|}.$$

By the hypothesis of Lemma 4, we know that $\left|\varphi'_{t(r)}(x)\right| > 0$. Therefore, to prove the desired conclusion (3), it suffices to prove that there exist constants $K < \infty$ and $\eta < 1$ such that for every tree $T \in \mathcal{F}_{\leq \boldsymbol{M}}$ with root r, all $L > 0$,

$$\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| \leq c_{t(r)} K \eta^{L-1}. \tag{5}$$

Note that K and η should be independent of T and L, but may depend on $\lambda, \varphi_1, \ldots, \varphi_t$ and $c_1, \ldots, c_t$. The constant K will be the following:

$$K := \frac{\lambda \Delta \max\limits_{1 \leq \ell \leq t} \sup\limits_{x \in I} |\varphi'_\ell(x)|}{\min\limits_{1 \leq \ell \leq t} c_\ell},$$

and the constant η will be the constant implicit in Condition $(\star\star)$.

We will show (5) by induction on L. First we verify the base case $L = 1$. In this case,

$$m_{L,r}(y) = \varphi_{t(r)}(\beta_{L,r}(y)) = \varphi_{t(r)}\left(\frac{1}{1 + \lambda y^{d(r)}}\right).$$

Thus,

$$\begin{aligned}
\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| &= \left|\frac{\partial \varphi_{t(r)}\left(\frac{1}{1+\lambda y^{d(r)}}\right)}{\partial y}\right| && \text{since } L = 1 \\
&\leq \sup_{x \in I}\left|\varphi'_{t(r)}(x)\right| \sup_{y \in [0,1]} \frac{\lambda d(r) y^{d(r)-1}}{\left(1 + \lambda y^{d(r)}\right)^2} && \text{by the chain rule} \\
&\leq \sup_{x \in I}\left|\varphi'_{t(r)}(x)\right| \lambda d(r) \\
&\leq \sup_{x \in I}\left|\varphi'_{t(r)}(x)\right| \lambda \Delta \\
&\leq c_{t(r)} K && \text{by the definition of } K.
\end{aligned}$$

This completes the analysis of the base case.

Now we proceed toward establishing the necessary induction step using the inductive hypothesis. We have that

$$\begin{aligned}
\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| &= \left|\frac{\partial F_{t(r)}\left(m_{L-1,w_1}(y), \ldots, m_{L-1,w_{d(r)}}(y)\right)}{\partial y}\right| \\
&= \left|\sum_{i=1}^{d(r)} \frac{\partial F_{t(r)}\left(m_1, \ldots, m_{d(r)}\right)}{\partial m_i} \cdot \frac{\partial m_{L-1,w_i}(y)}{\partial y}\right| && \text{where } m_i := m_{L-1,w_i}(y) \\
&= \left|\sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \ldots, m_{d(r)}\right)}{\partial m_i} \cdot \frac{1}{c_{t(w_i)}} \frac{\partial m_{L-1,w_i}(y)}{\partial y}\right|
\end{aligned}$$

$$= \left| \sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \ldots, m_{d(r)}\right)}{\partial m_i} \right|$$
$$\times \max_{1 \le i \le d(r)} \frac{1}{c_{t(w_i)}} \left| \frac{\partial m_{L-1,w_i}(y)}{\partial y} \right| \qquad \text{by Hölder's inequality.} \tag{6}$$

From $(\star\star)$, there exists a universal constant $\eta < 1$ such that

$$\left| \sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \ldots, m_{d(r)}\right)}{\partial m_i} \right| < \eta\, c_{t(r)}.$$

Therefore, it follows that

$$\begin{aligned} \left| \frac{\partial m_{L,r}(y)}{\partial y} \right| &\le \eta\, c_{t(r)} \cdot \max_{1 \le i \le d(r)} \frac{1}{c_{t(w_i)}} \left| \frac{\partial m_{L-1,w_i}(y)}{\partial y} \right| && \text{by (6) and the definition of } \eta \\ &\le c_{t(r)} K \eta^{L-1} && \text{by the inductive hypothesis.} \end{aligned}$$

This completes the proof of (5), and hence that of Lemma 4. □

4.1 Reproving Weitz's Result of SSM for Trees

In this section, we aim at finding a good choice of statistics. First we find such a statistic for the case $\boldsymbol{M} = [\Delta]$, i.e., the case of a single type, which enables us to reprove Weitz's result [34] that when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ SSM holds for every tree of maximum degree Δ.

Using Lemma 4 (and the simpler condition $(\star)$ for the case of a single type) we obtain a simpler proof of Weitz's result [34] that for every tree T with maximum degree $\Delta + 1$ (hence, for every graph G of maximum degree $\Delta + 1$) and for all $\lambda < \lambda_c(\mathbb{T}_{\Delta+1}) = \Delta^\Delta/(\Delta-1)^{\Delta+1}$, SSM holds on T (and on G).

Theorem 5. *Let $\varphi(x) = \frac{1}{s}\log\left(\frac{x}{s-x}\right)$ where $s = \frac{\Delta+1}{\Delta}$. Then, Condition $(\star)$ holds for $\boldsymbol{M} = [\Delta]$ and $\lambda < \lambda_c(\mathbb{T}_{\Delta+1})$. Consequently, SSM and the conclusions of Theorem 3 hold for $\boldsymbol{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ and $\lambda < \lambda_c(\mathbb{T}_{\Delta+1})$.*

Proof. First, a straightforward calculation implies that

$$\left| \frac{\partial F}{\partial m_i} \right| = \frac{1-\alpha}{s-\alpha}(s - \alpha_i),$$

where $\alpha_i = \varphi^{-1}(m_i)$ and $\alpha = \left(1 + \lambda \prod_{i=1}^{\Delta} \alpha_i\right)^{-1}$.

Hence, we have

$$\begin{aligned} \|\nabla F\|_1 &= \sum_{i=1}^{\Delta} \left| \frac{\partial F}{\partial m_i} \right| \\ &= \sum_{i=1}^{\Delta} \frac{1-\alpha}{s-\alpha}(s - \alpha_i) \end{aligned}$$

$$\leq \frac{1-\alpha}{s-\alpha}\,\Delta\,\left(s - \left(\prod_{i=1}^{\Delta}\alpha_i\right)^{1/\Delta}\right) \qquad \text{by the arithmetic-geometric mean inequality} \tag{7}$$

$$= \frac{1-\alpha}{s-\alpha}\,\Delta\,\left(s - \left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta}\right). \tag{8}$$

We now use the following technical lemma.

Lemma 6.

$$\max_{x\in[0,1]} \frac{(1-x)\left(1+\frac{1}{\Delta}-(\frac{1-x}{\lambda x})^{\frac{1}{\Delta}}\right)}{1+\frac{1}{\Delta}-x} \;\leq\; \frac{\omega}{1+\omega},$$

where Δ is a positive integer and ω is the unique solution to $\omega(1+\omega)^{\Delta}=\lambda$.

Using the above inequality (8) with Lemma 6, we have that:

$$\|\nabla F\|_1 < 1 \quad \text{if} \quad \frac{\omega}{1+\omega}\cdot\Delta \;<\; 1,$$

where ω is the unique solution of $\omega(1+\omega)^{\Delta}=\lambda$. This leads to the desired condition $\lambda < \lambda_c(\mathbb{T}_{\Delta+1}) = \Delta^{\Delta}/(\Delta-1)^{\Delta+1}$ so that SSM holds for $\boldsymbol{M}=[\Delta]$. As we noted in Remark 1, this is equivalent to SSM for $\boldsymbol{M}=\begin{pmatrix}0 & \Delta+1\\ 0 & \Delta\end{pmatrix}$. This completes the proof of Theorem 5. □

Proof of Lemma 6. Let $\Phi_\Delta(x)=(\frac{1-x}{\lambda x})^{\frac{1}{\Delta}}$ and $f(x)=\frac{(1-x)(1+\frac{1}{\Delta}-\Phi_\Delta(x))}{1+\frac{1}{\Delta}-x}$. Since $\Phi'_\Delta(x)=-\frac{\Phi_\Delta(x)}{\Delta x(1-x)}$, Φ_Δ is a decreasing function in $[0,1]$ such that $\Phi_\Delta(0)=+\infty$ and $\Phi_\Delta(1)=0$. Therefore it has a unique fixed point that can be shown to be $\bar{x}=\frac{1}{1+\omega}$. Moreover, it is the case that $\Phi_\Delta(x)>x$ if and only if $x<\bar{x}$. To prove Lemma 6, we notice that $f'(x)=\frac{(1+\frac{1}{\Delta})(\Phi(x)-x)}{\Delta x(1+\frac{1}{\Delta}-x)^2}$, hence $f'(x)>0$ for $x<\bar{x}$ and $f'(x)<0$ for $x>\bar{x}$. This implies that f has a maximum at $\bar{x}$, namely $f(\bar{x})=\frac{\omega}{1+\omega}$. □

4.2 DMS Condition: A Sufficient Criterion

Theorem 5 suggests choosing $\varphi_j(x)=\frac{1}{s_j}\log\left(\frac{x}{s_j-x}\right)$ with appropriate parameters s_j for a general branching matrix $\boldsymbol{M}$. Under this choice, we obtain the following condition for SSM.

Definition 7 (DMS Condition)**.** *Given a $t\times t$ branching matrix $\boldsymbol{M}$ and $\lambda^*>0$, for $s_1,\dots,s_t>1$ and $\boldsymbol{c}=(c_1,\dots,c_t)>0$, let $\boldsymbol{D}$ and $\boldsymbol{S}$ be the diagonal matrices defined as*

$$D_{jj} = \sup_{\alpha\in\left[\frac{1}{1+\lambda^*},1\right]} \frac{(1-\alpha)\left(1-\theta_j\left(\frac{1-\alpha}{\lambda^*\alpha}\right)^{1/\Delta_j}\right)}{s_j-\alpha} \qquad \textit{and} \qquad S_{jj}=s_j,$$

where

$$\theta_j := \frac{\left(\prod_\ell c_\ell^{M_{j\ell}}\right)^{1/\Delta_j}}{\sum_\ell c_\ell s_\ell M_{j\ell}/\Delta_j} \qquad \textit{and} \qquad \Delta_j=\sum_\ell M_{j\ell}.$$

We say the DMS Condition holds for $\boldsymbol{M}$ and λ^* *if there exist $s_1,\dots,s_t>1$ and $\boldsymbol{c}>0$ such that:*

$$(\boldsymbol{DMS})\,\boldsymbol{c}<\boldsymbol{c}.$$

Theorem 7. *If the DMS Condition holds for* $\boldsymbol{M}$ *and* $\lambda^* > 0$*, then Condition* $(\star\star)$ *holds with the choice of* $\varphi_j(x) = \frac{1}{s_j}\log\left(\frac{x}{s_j - x}\right)$ *for all* $\lambda \le \lambda^*$*. Consequently, SSM and the conclusions of Theorem 3 hold for* $\boldsymbol{M}$ *and all* $\lambda \le \lambda^*$*.*

Proof. First, one can check that

$$\left|\frac{\partial F_j}{\partial m_i}\right| = \frac{1-\alpha}{s_j - \alpha}(s_{j_i} - \alpha_i),$$

where $\alpha_i = \varphi_{j_i}^{-1}(m_i)$ and $\alpha = \frac{1}{1+\lambda\prod_{i=1}^{\Delta_j}\alpha_i}$.

Hence, it follows that

$$\begin{aligned}
\sum_{i=1}^{\Delta_j} c_{j_i}\left|\frac{\partial F_j}{\partial m_j}\right| &= \frac{1-\alpha}{s_j-\alpha}\sum_{i=1}^{\Delta_j} c_{j_i}(s_{j_i} - \alpha_i) \\
&\le \frac{1-\alpha}{s_j-\alpha}\left(\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} - \Delta_j\left(\prod_{i=1}^{\Delta_j} c_{j_i}\alpha_i\right)^{1/\Delta_j}\right) \quad \text{by the arithmetic-geometric mean ineq.} \\
&= \frac{1-\alpha}{s_j-\alpha}\left(\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} - \Delta_j\left(\prod_{i=1}^{\Delta_j} c_{j_i}\right)^{1/\Delta_j}\left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right) \\
&= \frac{1-\alpha}{s_j-\alpha}\left(1 - \theta_j\left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right)\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} \quad \text{by the definition of } \theta_j \\
&\le \frac{1-\alpha}{s_j-\alpha}\left(1 - \theta_j\left(\frac{1-\alpha}{\lambda^*\alpha}\right)^{1/\Delta_j}\right)\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} \\
&\le D_{jj}\sum_{\ell} M_{j\ell}c_\ell s_\ell \quad \text{by the definition of } D_{jj} \\
&< c_j \quad \text{by the DMS condition.}
\end{aligned}$$

which satisfies the desired condition $(\star\star)$ of Lemma 4. This completes the proof of Theorem 7. □

5 Application to $\mathbb{Z}^2$ in the hard-core model

In this section, we show how to apply Theorem 7 and Theorem 3 to the two-dimensional integer lattice $\mathbb{Z}^2$ and improve the lower bound on $\lambda_c(\mathbb{Z}^2)$, resulting in the following theorem.

Theorem 8. *There exists a* $t \times t$ *matrix* $\boldsymbol{M}$ *such that* $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\le \boldsymbol{M}}$ *and the DMS Condition holds for* $\lambda^* = 2.3882$*.*

Therefore, the following hold for $\mathbb{Z}^2$ *for all* $\lambda \le \lambda^*$*:*

1. *SSM holds on* $\mathbb{Z}^2$*.*
2. *There is a unique infinite-volume Gibbs measure on* $\mathbb{Z}^2$*.*

3. *If $\boldsymbol{M}$ has maximum degree Δ, if $t = O(1)$ and $\Delta = O(1)$, then for every finite subgraph G of $\mathbb{Z}^2$, Weitz's algorithm [34] gives an FPAS for approximating the partition function $Z(G)$.*

4. *For every finite subgraph G of $\mathbb{Z}^2$, the Glauber dynamics has $O(n \log n)$ mixing time.*

We first illustrate our approach by showing that Theorem 8 holds with $\lambda^* = 1.8801$ for a simple choice of $\boldsymbol{M}$. We then explain how to extend the approach to higher λ.

The graph $\mathbb{Z}^2$ is translation-invariant, hence the tree $T_{\text{saw}}(\mathbb{Z}^2, v)$ is identical for every vertex $v \in \mathbb{Z}^2$. Fix a vertex, call it the origin $\boldsymbol{o}$, and let us consider $T_{\text{saw}}(\mathbb{Z}^2) = T_{\text{saw}}(\mathbb{Z}^2, \emptyset)$. Each path from the root of $T_{\text{saw}}(\mathbb{Z}^2)$ corresponds to a self-avoiding walk in $\mathbb{Z}^2$ starting at the origin. Any walk on $\mathbb{Z}^2$ starting at the origin $\boldsymbol{o}$ can be encoded as a string over the alphabet $\{N, E, S, W\}$ corresponding to North, East, South and West. The tree $T_{\text{saw}}(\mathbb{Z}^2)$ contains such strings, truncated the first time the corresponding walk completes a cycle. A relaxed notion of such a tree would be to truncate a walk only when a 4-cycle is completed. Denote such a tree by T_4, and clearly we have that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of T_4. Our first idea is to define a branching matrix $\boldsymbol{N}$ so that $T_4 \in \mathcal{F}_{\leq \boldsymbol{N}}$, and hence $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\leq \boldsymbol{N}}$.

To avoid cycles of length four, it is enough to track the last three steps of the walks. Labeling the paths using $\{N, E, S, W\}$ as mentioned above, their branching rule is easily determined. For example, a path labeled NWW is followed by paths labeled WWS, WWN and WWW which corresponds to adding the directions S, N and W respectively. As another example, a path labeled NWS is followed by paths labeled WSW and WSS corresponding to adding the directions W and S to the path, while adding the direction E would have resulted in a cycle of length 4. The number of types in the corresponding branching matrix is $\leq 4 + 4^2 + 4^3 \leq 5^3$. Indeed, we can reduce the representation of such paths by using isomorphisms between the generating rules among them. This results in 4 types in the following branching matrix $\boldsymbol{N}$:

$$\boldsymbol{N} = \begin{pmatrix} 0 & 4 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \tag{9}$$

where the type $i = 0, ..., 3$ of a vertex (walk) in the tree represents the fact that a continuation with a minimum of $4 - i$ additional edges are needed to complete a cycle of length 4.

See Figure 2 for an illustration of this branching matrix $\boldsymbol{N}$. One can verify that this branching matrix captures, inter alia, the self-avoiding walk trees from $\mathbb{Z}^2$:

Observation 2. *For any finite subgraph $G = (V, E)$ of $\mathbb{Z}^2$ and $v \in V$, $T_{\text{saw}}(G, v) \in \mathcal{F}_{\leq \boldsymbol{N}}$.*

For this branching matrix, one can check that the (DMS) condition of Theorem 7 holds with $\lambda^* = 1.8801$, $\boldsymbol{S} = \mathbf{Diag}(1.040, 1.388, 1.353, 1.255)$ and $\boldsymbol{c} = (0.266037, 0.100891, 0.100115, 0.0973861)$. Checking the DMS Condition for a given choice of parameters would have been a straightforward task, were it not for the irrationality of the coefficients D_{jj}. However, one can establish rigorous upper bounds for D_{jj}, based on concavity of the function (of α) used in the definition of D_{jj}, in a suitable range of the parameters. These details will be discussed further below. As a consequence, we can conclude that Theorem 8 holds for $\boldsymbol{N}$ and $\lambda^* = 1.8801$.

The primary reason why the branching matrix $\boldsymbol{N}$ improves beyond the tree-threshold of $\lambda < \lambda_c(\mathbb{T}_4) = 27/16 = 1.6875$ is that the average branching factor of any $T \in \mathcal{F}_{\leq \boldsymbol{N}}$ is significantly smaller than that of the regular tree of degree 4.

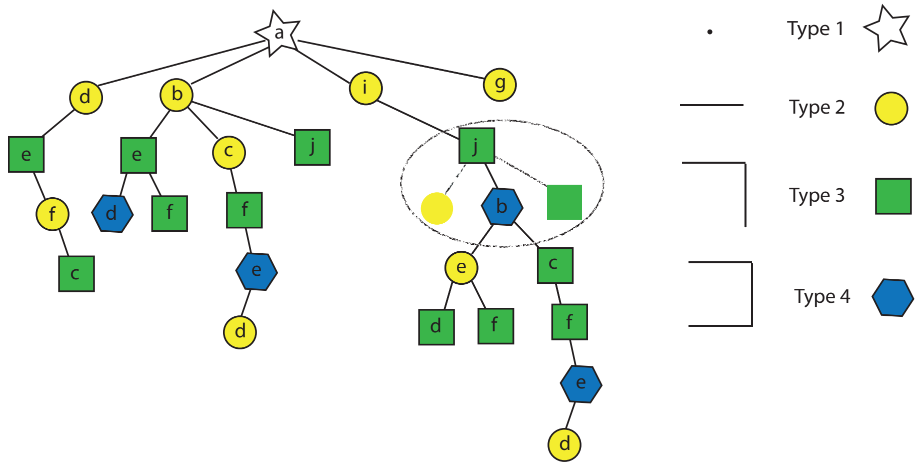

Figure 2: Assignment of the four types from matrix $\boldsymbol{N}$ defined in (9) to the self-avoiding walk tree T_{saw} from Figure 1. In the circled area, we also draw redundant leaves at vertex j which may appear in the branching rule, but not in T_{saw}.

To obtain a further reduction in the average branching, we observe that $\boldsymbol{N}$ did not consider the effect of occupying (or unoccupying) certain leaves as prescribed in Weitz's construction. Starting with T_4, prune the leaves as is done in the construction of $T_{\text{saw}}(\mathbb{Z}^2)$ from Section 2.2. Denote the new tree as T_4'. Clearly we still have that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of T_4'.

Let us illustrate the difference between T_4 and the pruned tree T_4'. We first fix an underlying order for the neighbors of each vertex. To this end, say $N > E > S > W$ and this prescribes an ordering of the neighbors of each vertex. Consider a leaf vertex v' in the tree T_4 corresponding to the vertex v in $\mathbb{Z}^2$ and to the path ρ in $\mathbb{Z}^2$. Since v' is a leaf vertex in T_4, ρ must end with a cycle at v, say $WNES$. Since v was exited in the West direction at the beginning of the 4-cycle, and since $W < N$, the leaf vertex v' would be labeled occupied in Weitz's construction, thus resulting in the removal of v' and its parent in the construction of T_4'. Note, every vertex w' in T_4 of type WNE has a child v' of type NES, and consequently w' (and its subtree) will be removed from the tree in the pruning process to construct T_4'. Thus, after removing vertices of type WNE (and similarly, WSE, SEN and ENW) from T_4, it is still the case that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of the resulting tree (T_4'). This highlights why T_4' has a significantly smaller average branching factor than T_4.

We can define a branching matrix $\boldsymbol{M}_2$, with 17 types (as illustrated in Figure 3), such that $T_4' \in \mathcal{F}_{\leq \boldsymbol{M}_2}$, and hence $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\leq \boldsymbol{M}_2}$. We can prove the DMS Condition is satisfied for $\boldsymbol{M}_2$ at $\lambda^* = 2.1625$, as we will describe shortly, which significantly improves upon our initial bound resulting from considering T_4.

A natural direction for improved results is to consider branching matrices corresponding to avoidance of larger cycles, while also accounting for the removal of vertices prescribed by the construction of Weitz. We briefly outline such an approach for walks avoiding cycles of length at most 4, 6, and 8, respectively. Avoiding cycles of length $2i$ results in $\sum_{j \leq 2i-1} 4^j \leq 5^{2i-1}$ types, hence the computations become increasingly difficult for larger i. For 8-cycles the task of finding appropriate parameters to satisfy the conditions of Theorem 7 is still feasible.

More precisely, we can define branching matrices $\boldsymbol{M}_i$ for $i \geq 2$, that (i) represent the structure of trees of walks avoiding cycles of length $\leq 2i$, as well as (ii) account for the removal of vertices

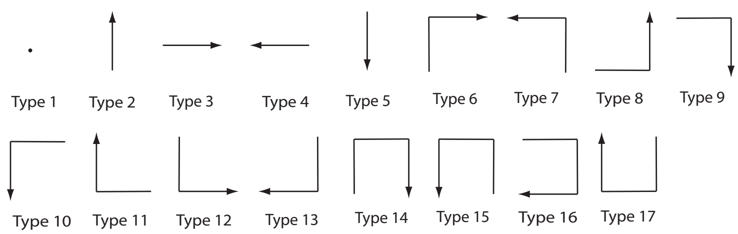

Figure 3: Shapes that the seventeen types (or labels) represent for $\boldsymbol{M}_2$ where $T_4' \in \mathcal{F}_{\leq \boldsymbol{M}_2}$.

based on children being labeled 'occupied.' One can extend the above construction of $\boldsymbol{M}_2$ for general $i > 2$ by using types encoded by longer paths with length at most $2i$ and ruling out the types that either contain a cycle of length at most $2i$ or whose children end up being labeled occupied. We can make the following general observation from our construction.

Observation 3. *For any finite subgraph $G = (V, E)$ of $\mathbb{Z}^2$ and $v \in V$, $T_{\text{saw}}(G, v) \in \mathcal{F}_{\leq \boldsymbol{M}_i}$ for any $i \geq 2$.*

As mentioned earlier, the matrix $\boldsymbol{M}_2$ constructed above consists of 17 types. An explicit description of it is shown in the Online Appendix [35], along with the associated parameters $\boldsymbol{S}$ and $\boldsymbol{c}$ for which one can check the DMS Condition for $\lambda^* = 2.1625$; this establishes Theorem 8 for $\boldsymbol{M}_2$ and $\lambda^* = 2.1625$.

The following table summarizes the threshold λ^* we obtain for each $\boldsymbol{M}_i$:

Max length of Avoiding-cycles	Effect of Occupations	Number of Types	λ^*
4	No	4	1.8801
4	Yes	17 ($< 5^3$)	2.1625
6	Yes	132 ($< 5^5$)	2.3335
8	Yes	922 ($< 5^7$)	2.3882

Note that, one can further improve the bound on λ by using more types for higher i and hence Theorem 8 on $\mathbb{Z}^2$ will hold with the corresponding activity λ^*. For any such matrix, the verification of the DMS Condition relies on (i) 'guessing' appropriate values for the parameters $\boldsymbol{S}$ and $\boldsymbol{c}$ and (ii) formally verifying that DMS Condition holds for the chosen $\boldsymbol{S}$ and $\boldsymbol{c}$. In choosing desirable $\boldsymbol{S}$ and $\boldsymbol{c}$, we employed a heuristic random walk algorithm.

To verify that the DMS Condition holds for a given *rational* matrix $\boldsymbol{S}$ and vector $\boldsymbol{c}$ is straightforward, provided we can obtain a rational upper bound for each type j for the function:

$$f_j(\alpha) = \frac{(1-\alpha)\left(1 - \theta_j \left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right)}{s_j - \alpha}.$$

Indeed, due to the concavity of this function for $0 < \theta_j \leq 1$, $s_j > 51/50$ and $\lambda > 27/16$, [2] it is always possible to find a *provable* upper bound for f_j in such a regime. This can be done, for

[2]This is a nontrivial algebraic fact. It can be proved by transforming the second derivatives condition to a set of integer polynomial constraints and using the "**resolve**" function in MATHEMATICA for the satisfiability of the constraints, which is rigorous by the Tarski-Seidenberg Theorem [32] for the real polynomial systems [36] and the so-called cylindrical algebraic decomposition [2].

example, by describing a suitable 'envelope' for f_j consisting of a piecewise linear function of the form:

$$g_j(\alpha) = \begin{cases} B_\ell & \text{if } \alpha < \alpha_\ell \\ \min\{b_\ell(\alpha - \alpha_\ell) + B_\ell, b_u(\alpha - \alpha_u) + B_u\} & \text{if } \alpha_\ell < \alpha < \alpha_u \\ B_u & \text{if } \alpha > \alpha_u \end{cases}$$

where α_ℓ, α_u are points such that $b_\ell > f_j'(\alpha_\ell) > 0$, $b_u < f_j'(\alpha_u) < 0$, $B_\ell > f_j(\alpha_\ell)$ and $B_u > f_j(\alpha_u)$. It is clear for any such function that $g_j(\alpha) > f_j(\alpha)$, thus we obtain a provable upper bound for f_j using g_j.

For every $\boldsymbol{M}_i$ in the above table, we provide $\boldsymbol{S}$ and $\boldsymbol{c}$, along with appropriate envelopes that lead to upper bounds $\widehat{D}_{jj}$ for the corresponding D_{jj}. Then we verify that the DMS Condition holds for the given values of λ by replacing D_{jj} with $\widehat{D}_{jj}$. For $i = 2, 3, 4$ these values ($\boldsymbol{M}$, $\boldsymbol{S}$, $\boldsymbol{c}$, α_ℓ and α_u) are given in the Online Appendix [35].

6 Ising Model

The approach taken here for the hard-core model can also be employed to address corresponding questions in the well-studied Ising model. The Ising model, with inverse temperature parameter $\beta \geq 0$, on a finite graph $G = (V, E)$ is the model associated with the Gibbs distribution μ on $\Omega = \{-1, +1\}^{|V|}$ such that for $\sigma = [\sigma_i] \in \Omega$,

$$\mu(\sigma) \;=\; \frac{1}{Z} \exp\Big(\beta \sum_{(i,j)\in E} \sigma_i \sigma_j\Big),$$

where the normalizing constant is the partition function: $Z = Z(G, \beta) := \sum_{\sigma\in\Omega} \exp\left(\beta \sum_{(i,j)\in E} \sigma_i\sigma_j\right)$. The notions of SSM, the self-avoiding walk tree representation, and branching trees defined for the hard-core model extend identically to the Ising model (or, for that matter, any other 2-spin model). Moreover, an analog of Lemma 4 also follows easily for Ising. Then, by the use of an appropriate statistic φ, the following simpler analog of the DMS Condition can be proved for the Ising model.

Theorem 9. *Given a $t\times t$ branching matrix M and $\beta^* > 0$, suppose there exists $\boldsymbol{c} = (c_1, \ldots, c_t) > 0$ such that*

$$\tanh(\beta^*)\boldsymbol{M}\boldsymbol{c} < \boldsymbol{c}, \tag{10}$$

then SSM and the conclusions of Theorem 3 hold for $\boldsymbol{M}$ and all $\beta \in [0, \beta^*]$.

Proof. First we note that Theorem 1 holds in general for all two spin models including the Ising model. Hence, Corollary 2 and Remark 1 are applicable to the Ising model as well. Further, observe that the proof of Theorem 3 (i.e., the implications of SSM) still hold for the Ising model. Consequently, we can prove Theorem 9 using similar notation and proof approach as was used for Theorem 7.

Given a tree $T \in \mathcal{F}_{\leq \boldsymbol{M}}$ and configuration $\boldsymbol{\rho}$, let us define again $\alpha = \alpha_r^{\boldsymbol{\rho}}(T, \beta)$ as the probability that the root r of T is minus-spinned. (Recall that in the hard-core model this was the probability that r was unoccupied.) If $w_1, \ldots, w_k$ are the children of r and $T_1, \ldots, T_k$ are the corresponding subtrees subtended at them, we let $\alpha_i := \alpha_{w_i}^{\boldsymbol{\rho}}(T_i, \beta)$ for $i \leq k$. For $i > k$, we define $\alpha_i :=$

1/2. Further let $\vartheta_i = \frac{1-\alpha_i}{\alpha_i}$, and $\vartheta = \frac{1-\alpha}{\alpha}$. Using these notations, a straightforward recursion calculation with the partition function leads to the following:

$$\vartheta = \prod_{i=1}^{\Delta_j} \frac{\exp(2\beta)\vartheta_i + 1}{\vartheta_i + \exp(2\beta)}, \tag{11}$$

where j is the type of r, and $\Delta_j = \sum_\ell M_{j\ell}$.

Motivated by (11), the function F_j (defined in Section 4 for the hard-core model) can be redefined for the Ising model as follows.

$$F_j(m_1, \dots, m_{\Delta_j}) := \varphi_j \left(\prod_{i=1}^{\Delta_j} \frac{\exp(2\beta)\varphi_{j_i}^{-1}(m_i) + 1}{\varphi_{j_i}^{-1}(m_i) + \exp(2\beta)} \right),$$

where j_i is the type of child w_i and φ_j is the statistic for a vertex of type j. Further, we define $m := \varphi_j(\vartheta)$ and $m_i := \varphi_{j_i}(\vartheta_i)$. It follows from (11) that $m = F_j(m_1, \dots, m_{\Delta_j})$. Then, one can prove the 'Ising version' of Lemma 4 with the interval $I = [\exp(-2\beta\Delta), \exp(2\beta\Delta)]$ using the same arguments as those in the proof of Lemma 4. Further, using the same arguments as in the proof of Theorem 7, with the choice of $\varphi_j(x) := \log(x)$, we have that

$$\frac{\partial F_j}{\partial m_i} = \frac{\vartheta_i \left(e^{4\beta} - 1\right)}{\left(e^{2\beta}\vartheta_i + 1\right)\left(e^{2\beta} + \vartheta_i\right)} \le \tanh(\beta),$$

from which the desired condition (10) follows easily. This completes the proof of Theorem 9. □

Using Theorem 9 with branching matrices $\boldsymbol{M}$ analogous to those we employed in Section 5 for the hard-core model, we can prove that SSM holds for the Ising model on $\mathbb{Z}^d$ for all $\beta < \beta^*$ as detailed in the following table:

Dimension	β^*
2	0.392190
3	0.214247
4	0.148045
5	0.113347

In comparison, applying Weitz's general technique to $\mathbb{Z}^2$ implies SSM for $\beta < .34657$.

We do not investigate the Ising model further because there are much stronger results known for this model. Onsager [24] established that $\beta_c(\mathbb{Z}^2) = \log(1+\sqrt{2}) \approx 0.440686$. And for general trees, Lyons [20, Theorem 2.1] established the critical point for uniqueness.

7 Acknowledgements

We are grateful to Karsten Schwan and the CERCS group at Georgia Tech for lending the support of their GPU machines so that we could conduct efficient parallel computations.

References

[1] S. E. Alm. Upper bounds for the connective constant of self-avoiding walks. *Combinatorics, Probability and Computing*, 4(2):115-136, 1993.

[2] D. S. Arnon, G. E. Collins, S. McCallum. Cylindrical Algebraic Decomposition I: The Basic Algorithm. *SIAM Journal on Computing*, 865-877, 1984.

[3] R. J. Baxter, I. G. Enting and S. K. Tsang. Hard-square lattice gas. *Journal of Statistical Physics*, 22(4):465-489, 1980.

[4] V. Beffara and H. Duminil-Copin. The self-dual point of the two-dimensional random cluster model is critical for $q \geq 1$. To appear in *Probability Theory and Related Fields*. Preprint available from the arXiv at: http://arxiv.org/abs/1006.5073

[5] J. van den Berg and A. Ermakov. A new lower bound for the critical probability of site percolation on the square lattice. *Random Structures and Algorithms*, 8(3):199-212, 1996.

[6] J. van den Berg and J. E. Steif. Percolation and the hard-core lattice gas model. *Stochastic Processes and their Applications*, 49(2):179–197, 1994.

[7] G. R. Brightwell, O. Häggström, and P. Winkler. Nonmonotonic behavior in hard-core and Widom-Rowlinson models. *Journal of Statistical Physics*, 94(3):415–435, 1999.

[8] F. Cesi. Quasi–factorization of the entropy and logarithmic Sobolev inequalities for Gibbs random fields. *Probability Theory and Related Fields*, 120(4):569–584, 2001.

[9] R. L. Dobrushin. The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. *Functional Analysis and its Applications*, 2(4):302-312. 1968.

[10] R. L. Dobrushin and S. B. Shlosman. Constructive unicity criterion. In *Statistical Mechanics and Dynamical Systems*, J. Fritz, A. Jaffe, and D. Szasz, eds., Birkhäuser, New York, pp. 347-370, 1985.

[11] M. E. Dyer, A. Sinclair, E. Vigoda, and D. Weitz. Mixing in time and space for lattice spin systems: A combinatorial view. *Random Structures and Algorithms*, 24(4):461-479, 2004.

[12] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang. Improved Inapproximability Results for Counting Independent Sets in the Hard-Core Model. Preprint, 2011. Preprint available from the arXiv at: http://arxiv.org/abs/1105.5131

[13] D. S. Gaunt and M. E. Fisher. Hard-Sphere Lattice Gases. I. Plane-Square Lattice. *Journal of Chemical Physics*, 43(8):2840-2863, 1965.

[14] H.-O. Georgii. *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin, 1988.

[15] L. Goldberg, R. Martin and M. Paterson. Strong Spatial Mixing for Lattice Graphs with Fewer Colours. *SIAM Journal on Computing*, 35(2):486-517, 2005.

[16] C. Greenhill. The complexity of counting colourings and independent sets in sparse graphs and hypergraphs. *Computational Complexity*, 9(1):52–72, 2000.

[17] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution distribution. *Theoretical Computer Science*, 43(2-3):169-186, 1986.

[18] F. P. Kelly. Loss Networks. *Annals of Applied Probability*, 1(3):319-378, 1991.

[19] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times.* American Mathematical Society, 2008.

[20] R. Lyons. The Ising model and percolation on trees and tree-like graphs. *Communications in Mathematical Physics*, 125(2):337-353, 1989.

[21] F. Martinelli. Lectures on Glauber dynamics for discrete spin models. *Lectures on Probability Theory and Statistics (Saint-Flour, 1997)*, Lecture Notes in Mathematics, 1717:93-191, 1998.

[22] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Communications in Mathematical Physics*, 161(3):447-486, 1994.

[23] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case. *Communications in Mathematical Physics*, 161(3):487-514, 1994.

[24] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Physics Review Letters*, 65(3-4):117-149, 1944.

[25] A. Pönitz and P. Tittman. Improved upper bounds for self-avoiding walks in $\mathbb{Z}^d$. *Electronic Journal of Combinatorics*, 7(1):R21, 2000.

[26] Z. Rácz. Phase boundary of Ising antiferromagnets near $H = H_c$ and $T = 0$: Results from hard-core lattice gas calculations. *Physical Review B*, 21(9):4012-4016, 1980.

[27] D. C. Radulescu. *A computer-assisted proof of uniqueness of phase for the hard-square lattice gas model in two dimensions.* Ph.D. dissertation, Rutgers University, New Brunswick, NJ, USA,1997.

[28] D. C. Radulescu and D. F. Styer. The Dobrushin-Shlosman phase uniqueness criterion and applications to hard squares. *Journal of Statistical Physics*, 49(1-2):281-295, 1987.

[29] D. Randall. Slow mixing of Glauber dynamics via topological obstructions. In *Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 870-879, 2006.

[30] A. Sly. Computational Transition at the Uniqueness Threshold. In *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 287-296, 2010.

[31] D. Štefankovič, S. Vempala, and E. Vigoda. Adaptive Simulated Annealing: A Near-optimal Connection between Sampling and Counting. *Journal of the ACM*, 56(3):1-36, 2009.

[32] A. Tarski. *A Decision Method for Elementary Algebra and Geometry.* 2nd edition, University of California Press, 1951.

[33] L. G. Valiant. The Complexity of Enumeration and Reliability Problems. *SIAM J. Computing*, 8(3):410-421, 1979.

[34] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, 140–149, 2006.

[35] Data is available from: http://www.cc.gatech.edu/∼vigoda/hardcore.html

[36] See: http://reference.wolfram.com/mathematica/tutorial/RealPolynomialSystems.html